MAQC and Other Efforts on Microarray Quality Control and Standardization

Leming Shi

National Center for Toxicological Research (NCTR)
U.S. Food and Drug Administration (FDA)
Jefferson, Arkansas 72079, U.S.A.

Leming.Shi@fda.hhs.gov

For discussion at the MAQC Project Meeting,,
FDA/CDER, Rockville, MD, May 2-3, 2005.

Views expressed in this presentation are those of the presenter and not necessarily those of the U.S. FDA.

SPECIAL SECTION

GENES IN ACTION

NEWS

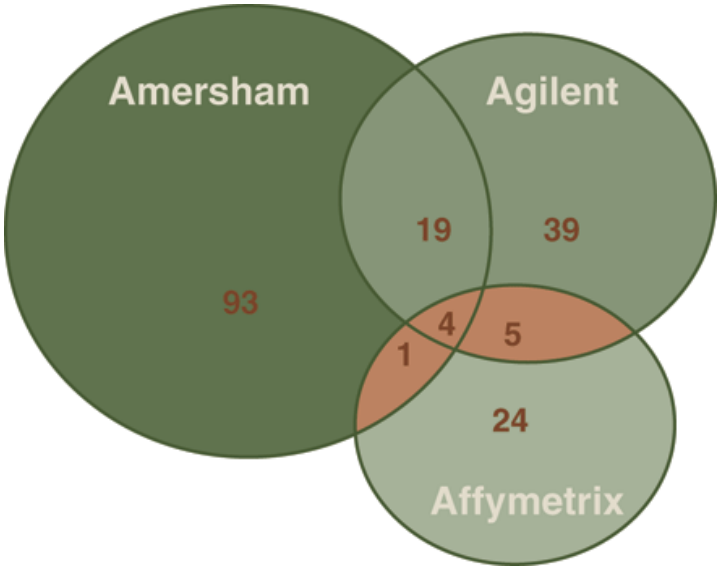
Getting the Noise Out of Gene Arrays

Thousands of papers have reported results obtained using gene arrays, which track the activity of multiple genes simultaneously. But are these results reproducible?

he gathered on kidney tumor cells, the less significant it seemed.

But those who have persevered with gene expression arrays attribute such problems to early growing pains. They claim

E. Marshall, *Science* 306, 630 (Oct 22, 2004).



"**Little overlap**."

"… the devices produced **wildly incompatible data**, largely because they were **measuring different things**."

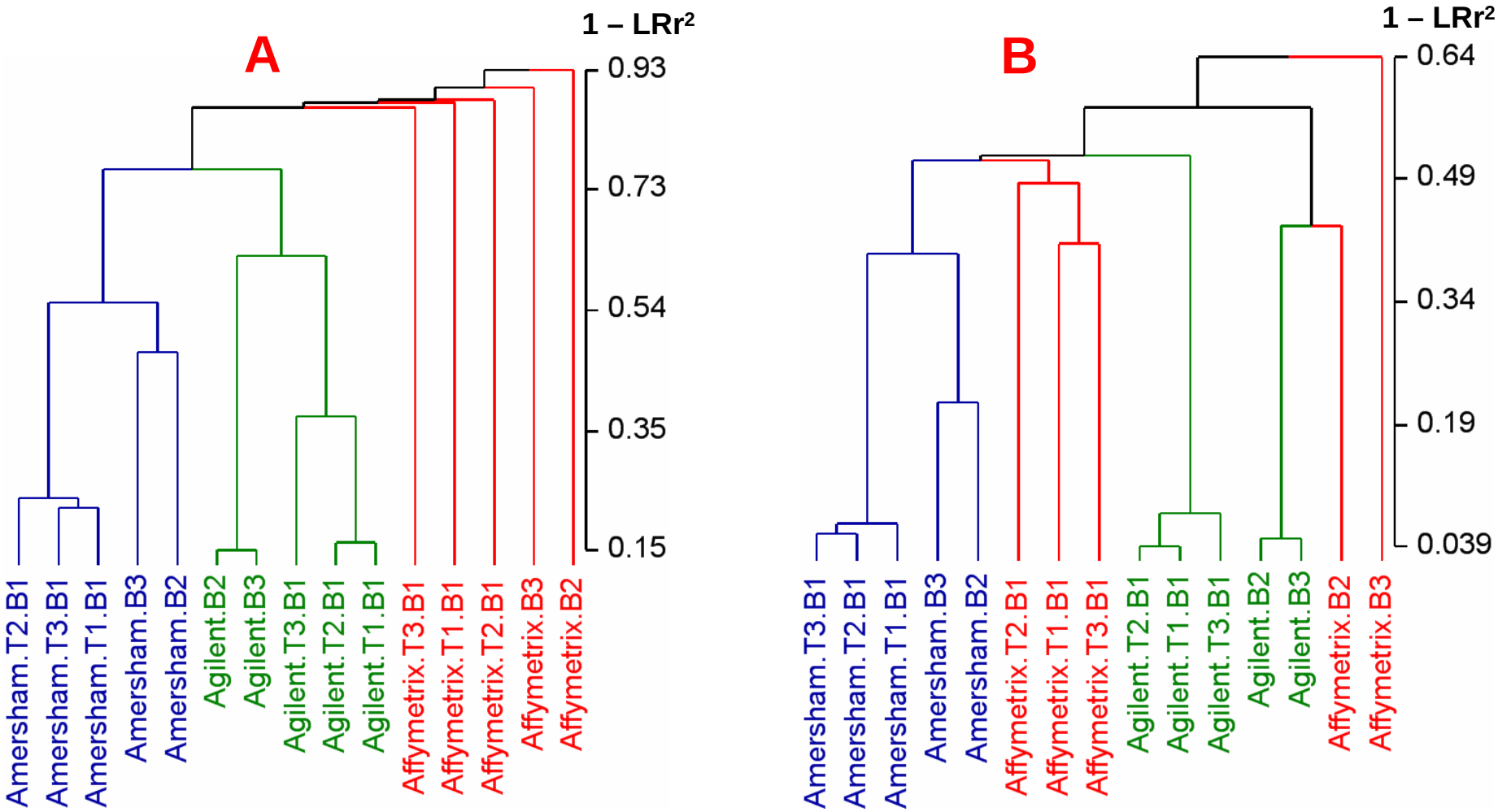
"… suggesting the need for establishing industrial **manufacturing standards**, and further independent and thorough validation of the technology."

P.K. Tan *et al.*, *Nucleic Acids Res* 31, 5676 (Oct 1, 2003).

???

- **Intra-platform/lab performance?**
- **Data analysis methods?**

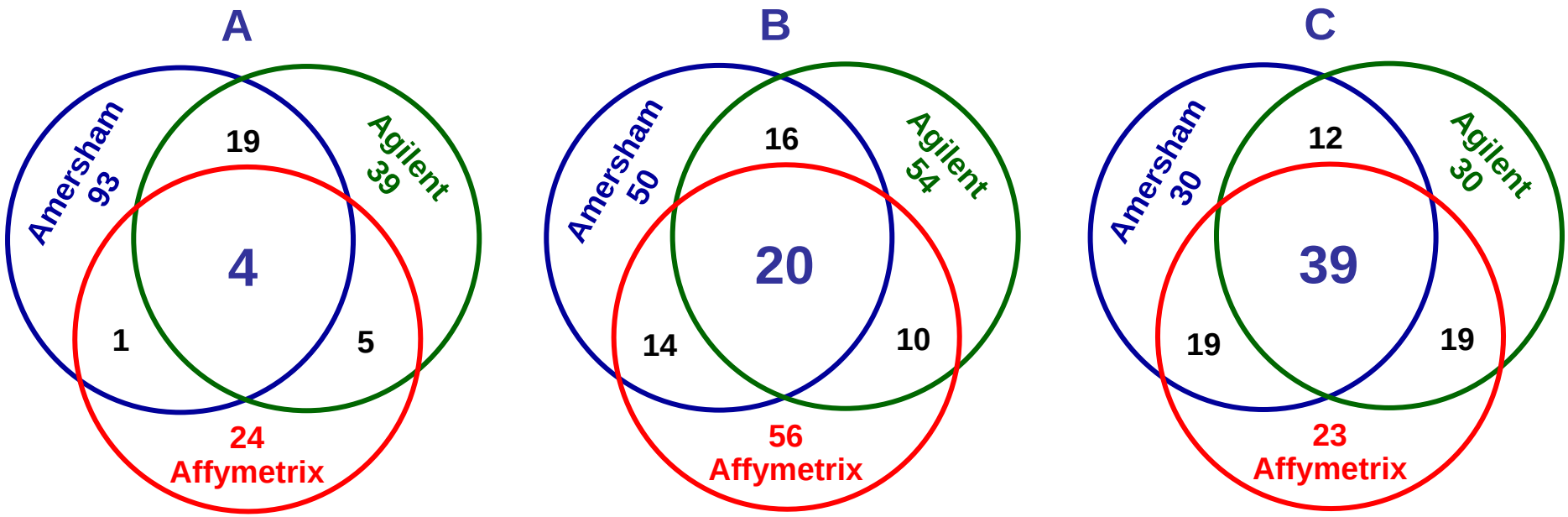
Poor Intra-platform Consistency



2009 genes → **Data (noise) filtering (ignoring 50% genes with the lowest average intensity on each platform)** → **537 genes**

Data filtering procedure: A. Barczak *et al.*, Genome Res 13, 1775 (2003)

Cross-platform Concordances Using Three Gene Selection Methods

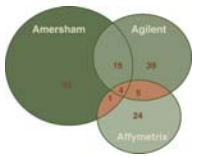


***p*-value cutoff** (*Tan et al 2003)*
(without noise filtering)

6 in common if 100 genes are selected from each platform.

SAM
(without noise filtering)

Fold-change ranking
(with noise filtering)



Shi L. *et al.*, *BMC Bioinformatics*, in press

Two Challenges
Facing the Microarray Community

- To ensure experimental proficiency of individual laboratories
- To objectively assess the merits of various data analysis methods

Because there is a lack of:

- Calibrated RNA **samples**
- Reliable benchmark **datasets**
- **Metrics/Thresholds** for assessing the performance achievable on microarray platforms
- Thorough and independent **validation**
- **Guidelines** for microarray QC and data analysis

Standardizing global gene expression analysis between laboratories and across platforms

Mouse liver RNA vs tissue mixture RNA (liver + kidney + lung + brain + spleen)

Members of the Toxicogenomics Research Consortium[1] NATURE METHODS | VOL.2 NO.5 | MAY 2005 | 351

Multiple-laboratory comparison of microarray platforms

Two mixture RNAs from 4 human knockout cell lines

Rafael A Irizarry[1], Daniel Warren[2], Forrest Spencer[3], Irene F Kim[4], Shyam Biswal[5], Bryan C Frank[6], Edward Gabrielson[7], Joe G N Garcia[8], Joel Geoghegan[9], Gregory Germino[4], Constance Griffin[10], Sara C Hilmer[11], Eric Hoffman[11], Anne E Jedlicka[12], Ernest Kawasaki[9], Francisco Martínez-Murillo[13], Laura Morsberger[10], Hannah Lee[5], David Petersen[9], John Quackenbush[6,14], Alan Scott[12], Michael Wilson[15,17], Yanqin Yang[2], Shui Qing Ye[8] & Wayne Yu[16]

NATURE METHODS | VOL.2 NO.5 | MAY 2005 | 345

The adoption of a common pair of readily accessible RNA samples will make such kind of studies much more valuable to the microarray community.

Experimental design, data analysis, and quality evaluation approaches to maximize cross-platform and cross-protocol inter-comparability of gene expression microarray data.

MGED 7, September 2004

Johannes Freudenberg[c], Sue Kong[c], Anil Jegga[c], Cathy Ebert[c], Shawn Smith[c], Craig Tomlinson[c], Maureen Sartor[c],Mario Medvedovic[c], Michael Wagner[c], Tinghu Qiu[n], Jeff Green[n], Shirley Shurtleff[j], James Downing[j], Anika Bissahoyo[u], Jennifer Clore[u], David Threadgill[u], Steve Settle[v], Braden Boone[v], Shawn Levy[v], Robert Coffey[v]; and Bruce Aronow[c]

Day one whole mouse RNA vs adult colon RNA

From the National Cancer Institute's Mouse Models of Human Cancer Consortium

The ERCC is producing standardized expression controls

- Well-characterized, widely accepted RNA standard controls for multiple platforms
  - Certified Reference Material (CRM)
- Protocols for multiple applications, research and clinical laboratory (CLSI/NCCLS)

1-color assays
- characterize the relationship between signal and RNA concentration

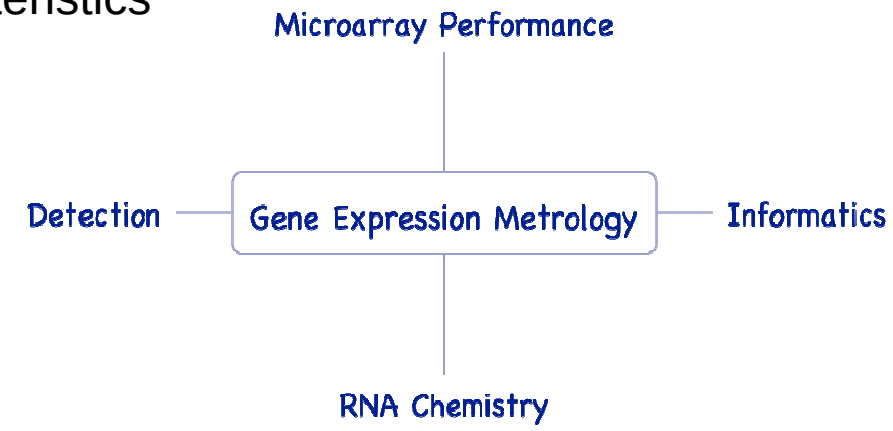
2-color assays
- detect known differences between two different spikes

QRT-PCR
- Assess C_t values

Courtesy of Dr. Janet Warrington

Elements of NIST's Gene Expression Metrology Program

- Microarray Performance
  - methodology for performance figures of merit
    - intrinsic performance, not comparative
  - uncertainty budgets
- Detection
  - validation strategies
  - spectroscopic characteristics
- Informatics
  - statistical metrology to underpin inference
- RNA Chemistry
  - hybridization thermodynamics and models
  - direct measures of mRNA quality



Courtesy of Marc Salit (NIST)

The ERCC, MAQC, and NIST Metrology Efforts Are Complementary

ERCC:

Provides QC indications for real sample hybridizations and array platforms

Does not guarantee good performance of "real genes"

MAQC:

Provides tools (RNAs, ref datasets, QC metrics/thresholds) for proficiency tests

Does not guarantee good quality of array data on real samples

NIST Gene Expression Metrology Program:

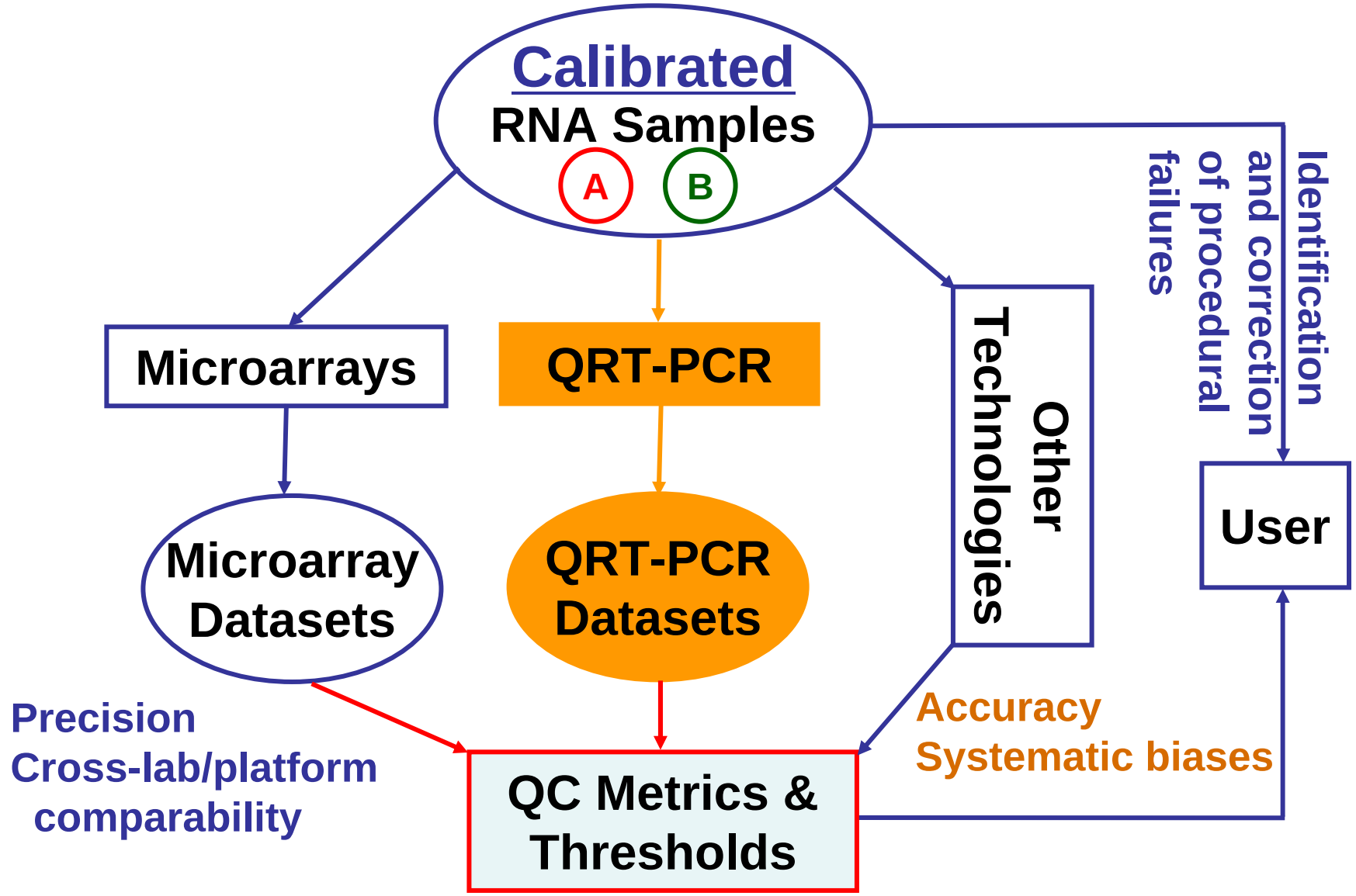
Provides better understandings on the fundamentals of microarray measurements

Addresses a wide range of issues

Coordination with ERCC and NIST's Metrology Program

1. Many shared participants and organizations.
2. Encourage the use of currently available spike-in controls.
3. Encourage the submission of in-house quality control metrics.
4. Recommend the use of reference RNA samples as backgrounds for testing ERCC spike-in controls.
5. Provide reference data for evaluating the performance of spike-in controls.
6. Share bioinformatics tools and QC metrics.

The MAQC Project: Microarray Quality Control



The MAQC Project:
A Community-wide Effort

- **A large effort involving many organizations including:**
  - Major microarray platform providers (Affymetrix, Agilent, Applied Biosystems, GE Healthcare, Illumina, and more to join …)
  - Major RNA sample providers (Ambion, Clontech, and Stratagene)
  - All FDA Centers (CBER, CDER, CDRH, CFSAN, CVM, NCTR)
  - Other organizations (EPA, NIST, Harvard, UMass, UCLA, ViaLogy…)
- **1st MAQC project meeting at FDA/NCTR, February 11, 2005**
- **MAQC Pilot Study: March-April, 2005**
- **2nd MAQC project meeting at FDA/CDER, May 2-3, 2005**
- **Complementary to and closely aligned with other efforts (e.g., ERCC, NIST Gene Expression Metrology Program)**
- **Everyone is invited to participate**
- **Results will be shared by the microarray community**

Selection of RNA Samples

Two RNA samples for each species (Human, Mouse, and Rat).

Starting with one species (Human).

Criteria for RNA sample selection:

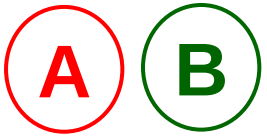
Available in large quantity
Reproducibility in production
High quality
Accessibility (commercial sources)
Wide gene presence
Large fold changes for a number of genes

Options for RNA sample selection:

1. Two universal reference RNAs
2. Two tissue-specific RNAs
3. Two cell lines
4. Combination

From Leming Shi, 1st MAQC Project Meeting, February 11, 2005

The MAQC Pilot Study



Four Candidate RNA Samples:

A. Ambion Brain RNA

B. Ambion Liver RNA

C. Clontech UHRR

D. Stratagene UHRR

UHRR: Universal Human Reference RNA

Clontech UHRR

"… is made by pooling the total RNA extracts from a collection of different human tissues,…"

Hundreds of mgs are currently available

www.clontech.com/clontech/archive/APR02UPD/pdf/ControlRNA.pdf

Stratagene Universal Reference RNA Preparation

A number of cell lines were selected from different tissues to gain optimal expression coverage for each species

CELL LINES

Human

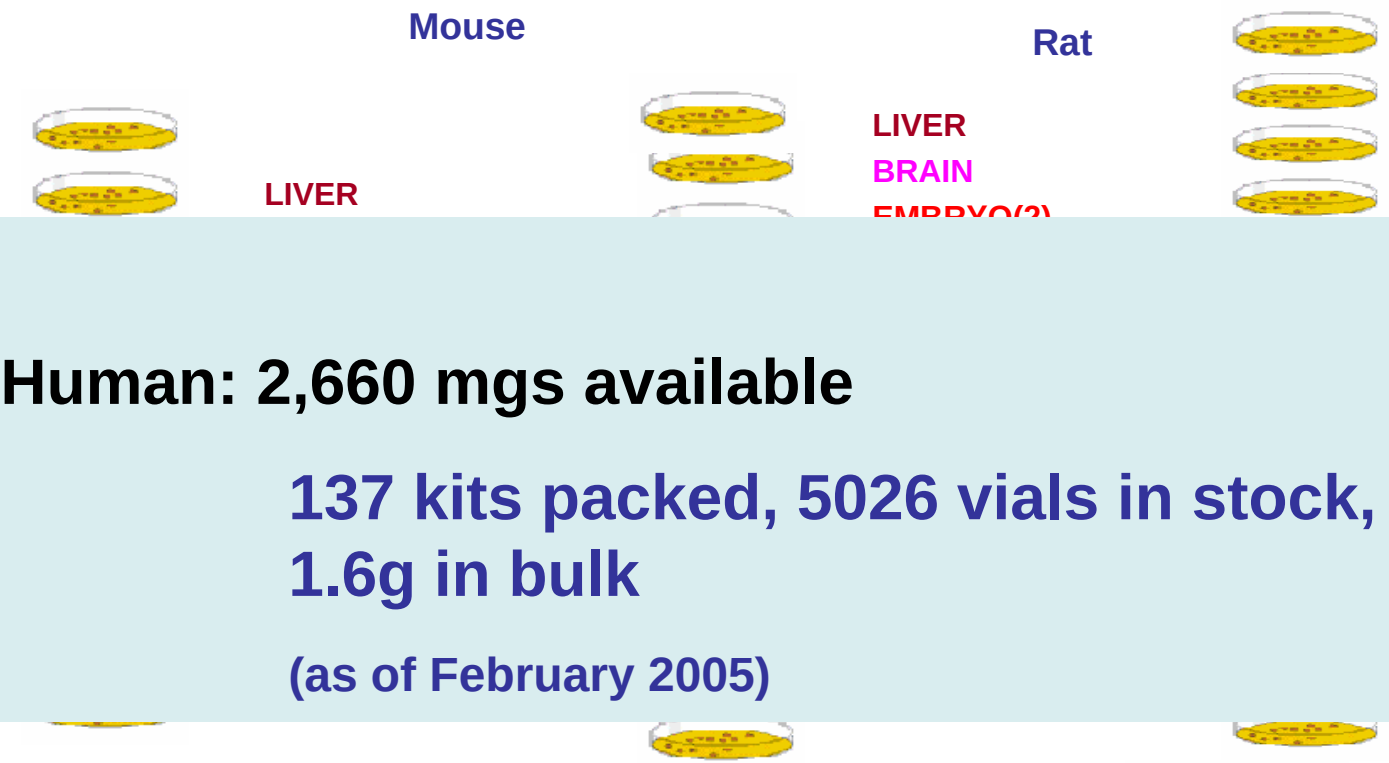
LIVER
LIPOSARCOMA
BRAIN
SKIN
BREAST
TESTIS
CERVIX
T-LYMPHOCYTE
B-LYMPHOCYTE
MACROPHAGES

Mouse

LIVER

Rat

LIVER
BRAIN
EMBRYO(2)

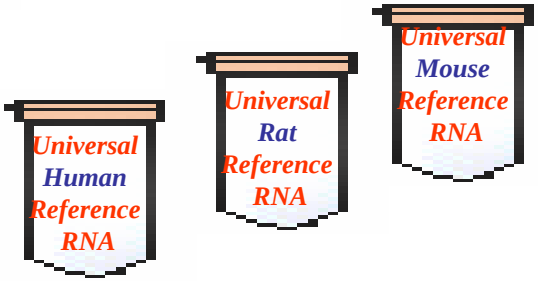


Human: 2,660 mgs available

137 kits packed, 5026 vials in stock, 1.6g in bulk

(as of February 2005)

RNA ISOLATION: equal quantities of total RNA from each cell line were pooled together



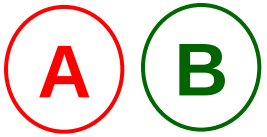
Courtesy of Dr. Gavin Fischer (Stratagene)

Ambion Brain RNA

Ambion Liver RNA

Invite Ambion's Mike Wilson / Bob Setterquist to comment

The MAQC Pilot Study



Four Platforms:

1. Affymetrix
2. Agilent
3. GE Healthcare
4. Illumina

Six Test Sites (7 datasets):

1. Affymetrix (Affymetrix)
2. Agilent (Agilent)
3. Ambion (Affymetrix and GEHC)
4. Illumina (Illumina)
5. NCTR (Agilent)
6. UMass Boston (GEHC)

MAQC Pilot Datasets

Five (5) replicates per sample per test site for one-channel platforms, resulting in **20** hybridizations per site per platform.

For Agilent platform, 6 sample-pairs were hybridized in 5 replicates, resulting in **30** hybridizations per test site (Agilent and NCTR).

160 hybridizations

MAQC Data Centralization/Distribution With ArrayTrack @ FDA/NCTR

April 5, 2005: Data submitted to NCTR

April 13, 2005: Data centralized within ArrayTrack and distributed to 11 sites agreed by the MAQC group for performing data analysis:

1. Affymetrix
2. Agilent
3. Ambion
4. Applied Biosystems
5. Clontech
6. GE Healthcare
7. Illumina
8. NCTR
9. NIST
10. Stratagene
11. UMass Boston

- RNA sample providers
- Platform providers
- Pilot Study test sites
- NIST

MAQC Project Planning Committee

http://edkb.fda.gov/webstart/arraytrack/

Terms and Conditions for Accessing the MAQC Pilot Study Datasets

April 12, 2005

Dear MAQC Pilot Study Data Analysis Site:

As we're preparing to distribute 7 datasets from the MAQC pilot study to each data analysis site, we would like to make sure that each site understands the following, as we discussed repeatedly during previous MAQC teleconferences:

1. Eleven (11) organizations (Affymetrix, Agilent, Ambion, Applied Biosystems, Clontech, FDA/NCTR, GE Healthcare, Illumina, NIST, Stratagene, and UMass Boston) will have full access to the 7 datasets, and each organization agrees to **conduct its independent analysis of the 7 datasets with its own preferred procedures** in order to **rank the 4 candidate RNA samples by gene coverage** (Table 1, attached) and to **rank the 6 sample pairs by ratio dynamic range** (Table 2, attached).
2. Each site also agrees to **rank the 4 RNA samples by QC measurements** provided by the 6 test sites and RNA sample providers (Table 3, attached).
3. Each organization agrees to fill in the attached Excel spreadsheet (MAQCpiliot_RankingRNAs_Organization.xls) and submit the results to Leming Shi (Leming.Shi@fda.hhs.gov) by **April 29, 2005** (please rename the file by substituting "Organization" with the name of your organization). Each site should give a brief presentation (~15 mins) on its analysis at the MAQC meeting in FDA/CDER, May 2-3, 2005.
4. Each site agrees that the purpose of the MAQC pilot study is solely for ranking the RNA samples and sample pairs so that two RNA samples will be selected for the MAQC main study. Therefore, the datasets from the MAQC pilot study **should NOT be over-interpreted**, *e.g.*, for the assessment of platform performance and/or cross-platform comparability.
5. **No organization should disseminate the MAQC pilot study datasets to others.**
6. **Public presentation and/or publication of the MAQC pilot study results without the consent of the MAQC participants are prohibited.**

If you agree to these, please reply to this message. I'll then e-mail you with information for accessing the MAQC pilot study datasets and the RNA quality data.

Table 1:
Ranking RNA Samples By Gene Coverage

	Table 1. Ranking of 4 RNA Samples by Gene Expression (Presence)							
RNA Sample	Datasets:							Average ranking
	I	II	III	IV	V	VI	VII	
A. Ambion Brain	1	1	1	1	1	1	1	1.0
B. Ambion Liver	1	1	1	1	1	1	1	1.0
C. Clontech UHRR	1	1	1	1	1	1	1	1.0
D. Stratagene UHRR	1	1	1	1	1	1	1	1.0
Error-checking	4	4	4	4	4	4	4	

Notes: (1) The 4 RNA samples should be ranked from 1 (most favorable) to 4 (least favorable) based on each of the 7 datasets individually; (2) Replace "1" (black font) with your actual ranking (1, 2, 3, or 4); and (3) Do NOT edit the cells in red font; they will be calculated autimatically as the average ranking by organization (last column) and for error-checking purposes (last row values should be equal to **10** (1+2+3+4)).

Datasets:	I. Affymetrix_Affymetrix		IV. Agilent_NCTR		VII. Illumina_Illumina	
	II. Affymetrix_Ambion		V. GEHC_Ambion			
	III. Agilent_Agilent		VI. GEHC_UMass			

Table 2:
Ranking RNA Sample Pairs By Fold Change

	Table 2. Ranking of 6 RNA Sample Pairs by Fold Change							
RNA Sample Pair	Datasets:							Average ranking
	I	II	III	IV	V	VI	VII	
1. A-B	1	1	1	1	1	1	1	1.0
2. A-C	1	1	1	1	1	1	1	1.0
3. A-D	1	1	1	1	1	1	1	1.0
4. B-C	1	1	1	1	1	1	1	1.0
5. B-D	1	1	1	1	1	1	1	1.0
6. C-D	1	1	1	1	1	1	1	1.0
Error-checking	6	6	6	6	6	6	6	

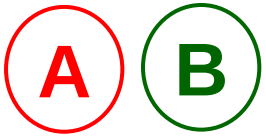
Notes: (1) The 6 RNA sample pairs should be ranked from 1 (most favorable) to 6 (least favorable) based on each of the 7 datasets individually; (2) Replace "1" (black font) with your actual ranking (1, 2, 3, or 4); and (3) Do NOT edit the cells in red font; they will be calculated autimatically as the average ranking by your organization (last column) and for error-checking purposes (last row values should be equal to **21** (1+2+3+4+5+6)).

Table 3:
Ranking Samples By RNA QC Data

	Table 3. Ranking of RNA Samples by RNA Quality Control Data							
RNA Sample	RNA QC data from:							Average
	Affymetrix	Agilent	Ambion	Illumina	NCTR	UMass	Provider	Ranking
A. Ambion Brain	1	1	1	1	1	1	1	**1.0**
B. Ambion Liver	1	1	1	1	1	1	1	**1.0**
C. Clontech UHRR	1	1	1	1	1	1	1	**1.0**
D. Stratagene UHRR	1	1	1	1	1	1	1	**1.0**
Error-checking	4	4	4	4	4	4	4	

Notes: (1) The 4 RNA samples should be ranked from 1 (most favorable) to 4 (least favorable) based on RNA QC data from each of the 6 test sites and the RNA providers individually; (2) Replace "1" (black font) with your actual ranking (1, 2, 3, or 4); and (3) Do NOT edit the cells in red font; they will be calculated autimatically as the average ranking by your organization (last column) and for error-checking purposes (last row values should be equal to **10** (1+2+3+4)).

Criteria for the Selection of RNA Samples



1. Available in large quantity
2. Reproducibility in production
3. High quality
4. Accessibility (commercial sources)
5. Wide gene presence
6. Large fold changes for a number of genes

From Leming Shi, 1st MAQC Project Meeting, February 11, 2005

Overall Ranking of RNA Samples

Proposed Formulae for the Overall Ranking of RNA Samples

	Criterion-1	Criterion-2	Criterion-3	Criterion-4	Criterion-5	Criterion-6	(Weighted?)
Sample	Available in large quantity	Reproducibility in production	Quality	Accessibility	Gene presence	Fold changes	Overall Ranking
A	2.5	2.5	2.5	2.5	2.5	2.5	2.5
B	2.5	2.5	2.5	2.5	2.5	2.5	2.5
C	2.5	2.5	2.5	2.5	2.5	2.5	2.5
D	2.5	2.5	2.5	2.5	2.5	2.5	2.5
Checking	10	10	10	10	10	10	
Weight	a	b	c	d	e	f	

Deconvolution of sample-pair based ranking (f.c.) to sample based ranking (Criterion-6):

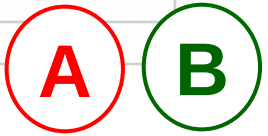
Sample	Pair-Ranking1	Pair-Ranking2	Pair-Ranking3	Average	Scaled Ranking
A	x (B)	x (C)	x (D)		
B	x (A)	x (C)	x (D)		
C	x (A)	x (B)	x (D)		
D	x (A)	x (B)	x (C)		

RNA Sample Pair
1. A-B
2. A-C
3. A-D
4. B-C
5. B-D
6. C-D

(Weighted) Overall Ranking =

a * Criterion-1 + b * Criterion-2 + c * Criterion-3 + d * Criterion-4 + e * Criterion-5 + f * Criterion-6

Excel Spreadsheet



2nd MAQC Project Meeting
May 2-3, 2005 @ FDA/CDER

- Select two RNA samples
- Design the MAQC Main Study (microarrays)
- Select 1,000 genes for QRT-PCR

MAQC Main Study (July-August, 2005):

~1000 hybridizations?

2nd MAQC Project Meeting Agenda

Day 1 8:00 AM – 5:00 PM, Monday, May 2, 2005

Morning

- MAQC and Other Efforts on Microarray Quality Control and Standardization
- Analysis of Datasets from the MAQC Pilot Study

12:00 pm Lunch (on your own)

Afternoon

- Decision on the Two RNA Samples
- Titration Strategies for Assessing the Quality of Microarrays
- Verification with Independent Platforms

Day 2 8:00 AM – 12:00 PM, Tuesday, May 3, 2005

MAQC Main Study

- Review of Slides Presented at the First MAQC Meeting on February 11, 2005
- Assessing Precision and Reproducibility
- Assessing Accuracy (Biases)
- Timeframe
- Rodents

12:00 pm Close of Meeting

Acknowledgments

FDA/NCTR

Toxicoinformatics
Megan Cao
Hong Fang
Steve Harris
Huixiao Hong
Roger Perkins
Feng Qian
Leming Shi
Zhenqian Su
Hongmei Sun
Weida Tong
Qian Xie

Biometry
Jim Chen

Functional Genomics
Jim Fuscoe
Tao Han

Systems Toxicology
Yvonne Dragan
Lei Guo

Neurotoxicology
Tucker Patterson

Bill Slikker, Jr., Deputy Director, FDA/NCTR

Dan Casciano, Director, FDA/NCTR

FDA/CBER
Jing Han
Raj Puri

FDA/CDER
Felix Frueh
Federico Goodsaid

FDA/CDRH
Gene Pennello
Uwe Scherf

FDA/CFSAN
Tom Cebula
Scott Jackson
Joseph LeClerc

FDA/CVM
Heather Harbottle

Non-FDA Collaborators
many…

Confirmed MAQC Participants

FDA Centers: CBER, CDER, CDRH, CFSAN, CVM, and NCTR

NIST: Marc Salit, Walter Liggett, David Deuwer, Mary Satterfield

EPA: David Dix, Wenjun Bao, Hongzu Ren, Chris Corton

Thank you!

Microarray Platform Providers

Affymetrix:	Janet Warrington, Jacques Retief
Agilent:	Jim Collins
Applied Biosystems:	Lu Zhang, Jack Zhai
GE Healthcare :	Timothy Sendera, Richard Shippy
Illumina:	Shawn Baker

RNA Sample Providers

Ambion:	David Dorris, Bob Setterquist, Mike Wilson
Clontech:	Laurence Lamarcq, Dmitry Bochkariov
Stratagene:	Gavin Fisher, Natalia Novoradovskaya

Others

UCLA/Cedars-Sinai:	Charles Wang
GenoSpectra:	Yuling Luo, Yunqing Ma
Harvard/Children's Hospital:	Zoltan Szallasi
NIH/NCI:	Ernest Kawasaki
UMass (Boston):	Roderick Jensen, Michael Lombardi
ViaLogy:	Bud Bromley

……

DISCLAIMER

- The selection of particular RNA samples for the MAQC project does NOT necessary imply that such RNA samples are better than other products.

- The two RNA samples are to be selected for research purpose only (MAQC project).